



# Expert Panel: Near-field and Long-range Model Evaluation Criteria

12<sup>th</sup> Modeling Conference On Air Quality Modeling

Moderated by James Thurman

US EPA/OAQPS/AQAD

Air Quality Modeling Group



# Discussion points

- Discussion of model evaluation techniques for near-field and long-range transport
  - EPA Protocol for determining best performing model (Cox-Tikvart protocol) for regulatory applications in near-field
    - Advantages/disadvantages
  - Episodic vs. long-term field studies
  - Non-regulatory applications (e.g. risk assessments)
  - Long-range transport evaluation needs
  - Key features of model evaluation databases



# Cox-Tikvart method (in a nutshell)

- EPA's protocol for model evaluation
- Robust highest concentration (RHC) and absolute fractional bias are heart of methodology
  - RHC is a measure of the top end of the concentration distribution (usually highest 26 values)
  - 1, 3, and 24-hour model to monitor comparisons based on monitored RHC and modeled RHC
  - 1-hour: paired in space and meteorological conditions (scientific component)
  - 3, 24-hour: unpaired in space and time (operational component)
    - Max monitor RHC, max modeled RHC
- Composite Performance Measure (CPM) combines 1,3, and 24-hour average absolute fractional biases
  - Can weight each averaging period depending on application
- Relative model performance measured using model comparison measure (MCM)
  - Compares model CPM values to each other
- Evaluations include bootstrapping to determine statistical significance across evaluated models



## Panelists

- Bret Anderson (U.S. Forest Service)
- Mark Garrison (ERM)
- Erik Snyder (EPA Region 6)



# Charge Questions

1. As part of the model evaluation process for establishing preferred models and approving alternative models for regulatory applications, the *Guideline* recommends the use of the EPA Protocol for Determining Best Performing Model, i.e., the Cox-Tikvart method to judge model performance. Is the Cox-Tikvart method still appropriate for near-field regulatory applications?

What are the advantages and disadvantages of the Cox-Tikvart protocol? How can or should applications that do not fit the Cox-Tikvart paradigm (*i.e.*, episodic or short-term tracer studies) be evaluated?
2. What evaluation methods, other than Cox-Tikvart, may be appropriate for consideration by EPA in updating the *Guideline* or could be used now for non-regulatory applications, such as risk assessments, where spatial and temporal distributions may be more important?
3. What evaluation methods and tools are available and appropriate for long range transport applications? In comparing the model evaluation needs for near-field and long-range transport application, what are the metrics most important or relevant to each and why do they differ?
4. What are the key features of model evaluation data sets for near-field models and long-range transport models?